XII. *On the Theory of Probabilities.*

By George Boole, *F.R.S., Professor of Mathematics in Queen's College, Cork.*

Received June 19,—Read June 19, 1862.

This paper has for its object the investigation of the general analytical conditions of a Method for the solution of Questions in the Theory of Probabilities, which was proposed by me in a work entitled "An Investigation of the Laws of Thought" (London, Walton and Maberly, 1854).

The application of this method to particular problems has been illustrated in the work referred to, and yet more fully in a 'Memoir on the Combination of Testimonies and of Judgments' published in the Transactions of the Royal Society of Edinburgh (vol. xxi. Part 4). Some observations, too, on the general character of the solutions to which the method leads, founded upon induction from particular cases, were contained in the original treatise, and the outlines, still in some measure conjectural, of their general theory were given in an Appendix to the Memoir. But the complete development of that theory was attended with analytical difficulties which I have only lately succeeded in overcoming. It involves discussions relating to the properties of a certain functional determinant, and to the possible solutions of a system of algebraic equations of peculiar form—discussions which will, I trust, be thought to possess a value, as contributions to Mathematical Analysis, independent of their present application.

As concerns the nature of the problems to which the method is applicable, it may be stated that they are such that the numerical elements, both given and sought, are the probabilities of events or states of things the definitions of which, and the connexions of which, are capable of expression by logical propositions. There is ground for believing that all questions whatever involving probability are ultimately reducible to this general form. This point, however, I do not purpose to discuss here. It has been already in some degree considered in the Memoir referred to.

In order to explain more fully the necessity for the present investigation, it will be requisite to state the fundamental principles upon which the method in question rests. There are only two of them which can possibly afford matter for discussion.

1st. The expression in language of the data of a problem in the Theory of Probabilities is to a certain extent arbitrary, because it depends upon the extent of meaning of the primary simple terms employed to express the events the conceptions of which it involves. But the choice of simple terms is, if we consider it with respect to our absolute power of choice, arbitrary. Any complex combination of events can be contemplated as a single whole in thought, and expressed by a single term. The invention of new simple terms to express what was before expressed by a combination of terms is a normal phenomenon in the growth of language.

Now the first principle upon which the method rests is the following:—

Principle I.—The different forms which a problem may be made to assume by different elections with respect to the simple terms of its expression are mutually equivalent.

For instance, if the following data were given,

The probability of rain is p,
The probability of rain with snow is q,

the form which the problem would assume in a language in which there was no word for snow, but in which the combination of snow with rain was called sleet, would be

The probability of rain is p,
The probability of sleet is q,

with the added condition, expressed as a logical proposition, that sleet always implies rain. And this as a statement of the data would, it is affirmed, be equivalent to the former statement. If these were the data of an actual problem, the event of which the probability is sought would require similar translation.

I desire to guard here against a possible misapprehension. I have said that the choice of simple terms, if considered with respect to our power of choice, is arbitrary. I do not mean by this to affirm that the actual growth of language is arbitrary. We know that it is far otherwise. Unity of sensuous impression in the early stages of its growth, unity of thought in the latter, seems to govern the invention and introduction of simple terms. It has indeed been said that there is a λόγος in the constitution of things of which language in its varied forms is the human reflexion, but never without the inseparable human element of choice and voluntary power.

It is then affirmed that whatever the grounds of fitness or propriety (and the existence of such grounds is fully conceded) may be, which have governed the actual choice of the simple terms of language, those grounds have nothing whatever to do with the calculation of probability. This depends upon the *information* contained in the data, information supposed to be derived from actual experience, or at least to be of such a nature that experience might have furnished it.

The different forms in which a problem is capable of being expressed, though differing in consequence of the different arbitrary elections which are possible with respect to its simple terms, are not independent of each other. They are connected together by the Laws of Thought, and pass one into the other by the processes of the Calculus of Logic, which is an organized expression of those Laws.

Among these forms there is one which presents exclusive advantages. It is that in which those events, however originally expressed, the probabilities of which constitute the data, are assumed as the simple events of the problem, and expressed by logical symbols corresponding to the simple terms of ordinary language; the event of which the probability is sought being also expressed logically by means of the same symbols. The Calculus of Logic enables us to do this, determining at the same time in an *explicit* form, *i. e.* in a form capable of expression in ordinary language by definite logical pro-

positions, the connexion which exists among all the events in question—a connexion which in the original form of the data was only *implied.*

This leads us to the statement of the second Principle of the Method.

Principle II.—When the data have been translated into probabilities of events connected by conditions logical in form and explicitly known, the problem may be constructed from a scheme of corresponding ideal events which are free, and of which the probabilities are such that when they (the ideal events) are restricted by the same conditions as the events in the data, their calculated probabilities will become the same as the given probabilities of the events in the data.

To take a material illustration: the problem, in the form to which it is reduced by the Calculus of Logic in accordance with Principle I., might be represented by the supposition of an urn containing balls distinguished by certain properties, *e. g.* by colour, as white or not white, by form, as round or not round, by material, as ivory or not ivory, and by the supposition that, while these properties enter into every conceivable combination, all the balls in which certain combinations are found are attached by strings to the sides of the urn, so that only the balls in which the remaining combinations are realized can be drawn. Suppose, further, that the probabilities of drawing under the actual conditions a white ball, a round ball, an ivory ball, &c. are given, and the probability of drawing a free ball fully defined with respect to the above elements of distinction is required. The principle affirmed is that we must proceed as if the balls were all free, and with probabilities such that the calculated probability of drawing any one of the balls which under the previous supposition are free, would be the same as under that supposition it is given to be.

Confining ourselves to the above material case, I remark, that the supposed mode of solution represents, 1st, a *possible* order of things; 2ndly, an order of things in which no preference is given to any one combination over any other which falls under the same category, or mode of thought. All the procedure of the theory of probabilities is founded upon the mental construction of the problem from some hypothesis, either, 1st, of events known to be independent; or, 2ndly, of events of the connexion of which we are totally ignorant; so that, upon the ground of this ignorance, we can again construct a scheme of alternatives all equally probable, and distinguished merely as favouring or not favouring the event of which the probability is sought. In doing this we are not at liberty to proceed arbitrarily. We are subject, first, to the formal Laws of Thought, which determine the possible conceivable combinations; secondly, to that principle, more easily conceived than explained, which has been differently expressed as the "principle of sufficient reason," the "principle of the equal distribution of knowledge or ignorance*," and the "principle of order." We do not know that the distribution of

* Knowledge and ignorance being in the theory of probabilities supplementary to each other, the equal distribution of the one implies that of the other.

I take this opportunity of explaining a passage in the 'Laws of Thought,' p. 370, relating to certain applications of the principle. Valid objection lies not against the principle itself, but against its applica-

properties in the actual urn is the same as it is conceived to be in the ideal urn of free balls, but the hypothesis that it is so, involves an equal distribution of our actual knowledge, and enables us to construct the problem from ultimate hypotheses which reduce it to a calculation of combinations.

I pass from the particular and material to the general problem. In the form to which this is brought by the Calculus of Logic, the probabilities are those of events of which certain combinations are, as a logical consequence of the original definitions of those events, impossible. It might, at first sight, appear that this establishes a fundamental difference between this problem and that of the urn, in which certain combinations are prevented from issuing by a material hindrance. In the one case the restriction appears as logically necessary, in the other as only actual.

Upon this I remark, that the data of the problem in its ultimate reduced form *might* result from the same kind of dependence as in the actual data; that they, in fact, *would* thus result if the mind of the observer were capable of contemplating, and were in a position to contemplate, each of the events in this ultimate translated form simply as a whole, and of recording, through an approximately infinite series of observations, what combinations of those wholes come into being, and what do not, in the actual universe. What appears as necessary in the translated data would now appear as actual—as a result of observation; what is impossible would be received as non-existent. The question is, then, whether the difference between the conception of what is impossible from involving a logical contradiction, and the conception of what in the actual constitution of things never exists, is of a kind to affect expectation. I do not hesitate to say that it is not. We are concerned with events in so far as they are capable of happening or not happening, of combining or not combining; but we are not concerned with the reasons in virtue of which they happen or do not happen, combine or do not combine. If we went beyond this, we should enter upon a metaphysical question to which I presume that no answer can, upon rational grounds, be given, viz. upon the question whether, when two things or events are in the actual constitution of things incapable of happening together, it would, if our knowledge were sufficiently extended, be found that the resulting conceptions of them were logically inconsistent.

I have but one further observation on Principle II. to make. It is that in the general problem we are not called upon to interpret the ideal events. The whole procedure is, like every other procedure of abstract thought, formal. We do not say that the ideal events exist, but that the events in the translated form of the actual problem are to be considered to have such relations with respect to happening or not happening as a certain system of ideal events would have if conceived first as free, and then subjected, without their freedom being otherwise affected, to relations formally agreeing with those to which the events in the translated problem are subject.

---

tion through arbitrary hypotheses, coupled with the assumption that any result thus obtained is necessarily *the* true one. The application of the principle employed in the text, and founded upon the general theorem of development in Logic, I hold to be *not* arbitrary.

We are now able to explain more clearly the nature of the analytical investigation which will follow. Let $p_1, p_2, \ldots p_n$ represent the probabilities given in the data. As these will in general not be the probabilities of unconnected events, they will be subject to other conditions than that of being positive proper fractions, viz. to other conditions beside

$$\left.\begin{array}{l} p_1 \geqq 0,\ p_2 \geqq 0 \ldots p_n \geqq 0 \\ p_1 \leqq 1,\ p_2 \leqq 1 \ldots p_n \leqq 1. \end{array}\right\} \quad \ldots\ldots\ldots\ldots \quad (1.)$$

Those other conditions will, as will hereafter be shown, be capable of expression by equations or inequations reducible to the general form

$$a_1 p_1 + a_2 p_2 \ldots + a_n p_n + a \gtreqless 0,$$

$a_1, a_2, \ldots a_n, a$ being numerical constants which differ for the different conditions in question. These, together with the former, may be termed the conditions of possible experience. When satisfied they indicate that the data *may* have, when not satisfied they indicate that the data *cannot* have, resulted from actual observation. On the other hand, the ideal events are regarded as independent, and their probabilities, which enter as auxiliary quantities into the process of solution, are subject to no other condition than that of being positive proper fractions. It is the general object of the analytical investigation to establish the two following conclusions, viz.,—

1st. The probabilities of the ideal independent events, as involved in the method under consideration, will in the process be determinable, without ambiguity, as positive proper fractions whenever the data satisfy the conditions of possible experience, and not otherwise.

And, as a consequence of the above,

2ndly. The probability determined by the method will have such a value as it consistently might have had if, instead of being calculated from the data, it had been determined by observation under the same experience as the data.

These conclusions rest upon the ground of certain analytical theorems relating to functional determinants, and to the possible solutions of simultaneous algebraic equations, which will be demonstrated in this paper. But, in order to explain the application of those theorems, it will be necessary to show, first, how the "conditions of possible experience" in problems in the Theory of Probabilities may be determined; secondly, what the analytical method in question for the solution of such problems is.

Determination of the Conditions of possible Experience.

The method for determining the conditions of possible experience given in the 'Laws of Thought,' chap. xix., may be advantageously replaced by the following one, which is taken from the 'Memoir on the Combination of Testimonies and of Judgments,' already referred to.

Let the events in the data be resolved into the ultimate possible alternatives which they involve, and let the unknown probabilities of these alternatives be represented by λ, μ, ν, &c.; then, as the probability of each event in the data is equal to the sum of the

probabilities of the alternatives which it involves, we shall have a system of equations connecting λ, μ, ν, &c. with $p_1, p_2, \ldots p_n$, the probabilities supposed given. Again, $\lambda, \mu, \nu \ldots$, as probabilities, are subject to the conditions

$$\lambda \geqq 0,\ \mu \geqq 0,\ \nu \geqq 0, \ldots \text{ \&c.},$$

and, as alternatives mutually excluding each other, to the condition

$$\lambda+\mu+\nu+\ldots=1,$$

or the condition

$$\lambda+\mu+\nu+\ldots \leqq 1,$$

according as the alternatives in question together make up certainty or not.

Thus we have a system consisting of equations and inequations from which λ, μ, ν, &c. must be eliminated. To effect this elimination we must determine as many of the quantities $\lambda, \mu, \nu \ldots$ as possible from the equations, substitute their values in the inequations, and then eliminate the remainder of the quantities $\lambda, \mu, \nu \ldots$ by means of the theorem that if we have simultaneously

$$\lambda \leqq a_1,\ \lambda \leqq a_2, \ldots \lambda \leqq a_m,$$
$$\lambda \geqq b_1,\ \lambda \geqq b_2, \ldots \lambda \geqq b_n,$$

then we have the system of conditions of which the type is

$$a_i \geqq b_j,$$

a_i representing any one of the set $a_1, a_2, \ldots a_m$, and b_j any one of the set $b_1, b_2, \ldots b_n$. Thus there are mn conditions in all.

This method is illustrated in the following problem, in the expression and solution of which it is to be noticed, that when in the Calculus of Logic an event is represented by x, the event which consists in its not happening is denoted by $1-x$, or for brevity by $\bar{x}$; that when two events are represented by x and y, their concurrence is denoted by xy, the happening of the first without the second by $x\bar{y}$, and so on.

Problem. Given that the probability of the concurrence of the events x and y is p, of the events y and z, q, and of the events z and x, r. Required the conditions to which p, q, and r must be subject in order that the above data may be consistent with a possible experience.

Resolving the events xy, yz, xz into the possible alternations out of which they are formed, let us write

$$\text{Prob. } xyz=\lambda,\ \text{Prob. } xy\bar{z}=\mu,\ \text{Prob. } x\bar{y}z=\nu,\ \text{Prob. } \bar{x}yz=\varrho.$$

Then we have the equations

$$\lambda+\mu=p,\ \lambda+\varrho=q,\ \lambda+\nu=r,$$

together with the inequations

$$\lambda \geqq 0,\ \mu \geqq 0,\ \nu \geqq 0,\ \varrho \geqq 0,$$
$$\lambda+\mu+\nu+\varrho \leqq 1.$$

From the equations we find

$$\mu=p-\lambda,\ \varrho=q-\lambda,\ \nu=r-\lambda,$$

which, substituted in the inequations, give

$$\lambda \geqq 0,\ p-\lambda \geqq 0,\ q-\lambda \geqq 0,\ r-\lambda \geqq 0,$$
$$p+q+r-2\lambda \leqq 1;$$

and it only remains to eliminate λ. Now from the above,

$$\lambda \overline{\overline{<}} p,\ \lambda \overline{\overline{<}} q,\ \lambda \overline{\overline{<}} r,\ \lambda \overline{\overline{>}} 0,\ \lambda \overline{\overline{>}} \frac{p+q+r-1}{2},$$

therefore

$$p \overline{\overline{>}} 0,\ q \overline{\overline{>}} 0,\ r \overline{\overline{>}} 0,$$

$$p \overline{\overline{>}} \frac{p+q+r-1}{2},\ q \overline{\overline{>}} \frac{p+q+r-1}{2},\ r \overline{\overline{>}} \frac{p+q+r-1}{2}.$$

The last three conditions are reducible to the simpler form,

$$p \overline{\overline{>}} q+r-1,\ q \overline{\overline{>}} r+p-1,\ r \overline{\overline{>}} p+q-1.$$

Such are the conditions of possible experience in the data.

Suppose, for instance, it was affirmed as a result of medical statistics that, in two-fifths of a number of cases of disease of a certain character, two symptoms x and y were observed; in two-thirds of the cases the symptoms y and z were observed; and in four-fifths of the cases the symptoms z and x were observed; so that, the number of cases observed being large, we might on a future outbreak of the disease consider the fractions $\frac{2}{5}$, $\frac{2}{3}$, and $\frac{4}{5}$ as the probabilities of recurrence of the particular combinations of the symptoms x, y, and z observed. The above formulæ would show that the evidence was contradictory. For, representing the respective fractions by p, q, and r, the condition $p \overline{\overline{>}} q+r-1$ is not satisfied. (*Edinburgh Memoir.*)

In applying the above method to the *à priori* limitation of questions in the theory of probabilities, it will be necessary to represent the probability sought by a single letter u, and treat this as if it were one of the numerical data. The resolution of the event of which the probability is sought into alternatives belonging to the same scheme as those of the events in the data gives us a new equation, which must be combined with the equations involving p, q, r, &c. The elimination of λ, μ, ν, &c. then determines not only the conditions of possible experience limiting p, q, r, but also the conditions which u must satisfy *à priori*, whatever method for its actual determination may be employed.

Thus, if from the foregoing data it were required to determine the *à priori* limits of Prob. xyz, i. e. of the probability of the conjunction of the events x, y, z, we should have as the additional equation

$$u=\lambda,$$

and therefore, after elimination of λ, μ, ν,

$$u \overline{\overline{<}} p,\ u \overline{\overline{<}} q,\ u \overline{\overline{<}} r,$$

$$u \overline{\overline{>}} 0,\ u \overline{\overline{>}} \frac{p+q+r-1}{2},$$

the conditions required.

It will, however, in most of the following investigations suffice to consider the conditions of possible experience in the data alone, because it will be shown that when these are satisfied the corresponding conditions for the probability sought, when its value is determined by the method of the following section, will also be satisfied.

Statement of the Method for the Solution of Questions in the Theory of Probabilities.

For the general demonstration of this method the reader is referred to the 'Laws of Thought,' chap. xvii. For the purpose of the analytical investigation the statement of the method will suffice.

Let s, t, v, &c. represent the events of which the probabilities are given, p, q, r, &c. those probabilities, and w the event of which the probability is sought; then, whatever the definitions of $s, t \ldots$ and w may be, and whatever connecting relations may exist, it is always possible by the Calculus of Logic to determine the logical dependence of w upon s, t, &c. in the following most general form, viz.

$$w = \mathrm{A} + 0\mathrm{B} + \frac{0}{0}\mathrm{C} + \frac{1}{0}\mathrm{D}.$$

Here A, B, C, D are logical combinations of the events s, t, &c., and the connexion in which these stand to the event w and to each other is the following: A expresses those combinations of s, t, &c. which are entirely included in w, i. e. which cannot happen without our being permitted to say that w happens. B represents those combinations which may happen but are not included under w; so that when they happen we may say that w does not happen. C represents those combinations the happening of which leaves us in doubt whether w happens or not. D those combinations the happening of which would involve logical contradiction.

It follows from the above that the *translated* form of the problem is

Given Prob. $s=p$, Prob. $t=q$, Prob. $v=r$, &c., s, t, $v \ldots$ being regarded as events subject to the explicit logical condition

$$\mathrm{A} + \mathrm{B} + \mathrm{C} = 1.$$

Required the probability u of the event of which the logical expression is

$$w = \mathrm{A} + \frac{0}{0}\mathrm{C};$$

and it is shown (Laws of Thought, p. 265), upon grounds essentially the same as those expressed in Principles I. and II. of this paper, that the solution of the problem is involved in the following *algebraic* equations, viz.

$$\frac{\mathrm{V}_s}{p} = \frac{\mathrm{V}_t}{q} \ldots = \frac{\mathrm{A} + c\mathrm{C}}{u} = \mathrm{V}, \quad \ldots \ldots \ldots \quad (\mathrm{I.})$$

in which the functions V, V_s, $\mathrm{V}_t \ldots$ are formed in the following manner, viz.,—

1st. V is derived from A+B+C without change of form by interpreting s, t, &c. no longer as logical symbols, but as symbols of quantity. They represent the probabilities of the ideal events of Principle II.

2ndly. V_s is the sum of those terms in V which contain s as a factor, V_t the sum of those which contain t as a factor, &c.

The quantity c is an arbitrary constant, admitting of any value between 0 and 1.

To effect the solution, the quantities s, t, &c. are to be eliminated from the system (I.), and u then determined as a function of p, q, $r \ldots$ and c. The arbitrary constant c may

not appear in the final result, because the developed form of w may not contain any terms affected with the symbol $\frac{0}{0}$. When such terms do appear, the constant c admits of an interpretation indicating what new data are required to make the solution definite*.

It is proper here to observe that the conditions of possible experience can be determined as well from the 'translated' as from the original form of the problem. That the results will agree is evident *à priori*, but it may be desirable to point out the analytical connexion of the two processes. I will take the example just considered, and then offer some general remarks on the subject.

Representing the events xy, yz, zx by s, t, v, the translated data would be found to be

$$\text{Prob. } s=p, \quad \text{Prob. } t=q, \quad \text{Prob. } v=r,$$

s, t, and v being connected by the explicit logical condition

$$stv+s\bar{t}\bar{v}+\bar{s}t\bar{v}+\bar{s}\bar{t}v+\bar{s}\bar{t}\bar{v}=1.$$

It is easily shown that the first member of this equation represents the sum of those combinations of the events s, t, v, with respect to happening or failing, which involve no logical contradiction.

If, then, we represent under the above condition

$$\text{Prob. } stv=\lambda', \quad \text{Prob. } s\bar{t}\bar{v}=\mu', \quad \text{Prob. } \bar{s}t\bar{v}=\nu', \quad \text{Prob. } s\bar{t}v=\varrho',$$

we shall have

$$\lambda'+\mu'=p, \ \lambda'+\nu'=q, \ \lambda'+\varrho'=r,$$
$$\lambda'\geqq 0, \ \mu'\geqq 0, \ \nu'\geqq 0, \ \varrho'\geqq 0,$$
$$\lambda'+\mu'+\nu'+\varrho'\leqq 1.$$

This system of equations and inequations agrees with that employed in the previous solution, if we only make

$$\lambda'=\lambda, \ \mu'=\mu, \ \nu'=\varrho, \ \varrho'=\nu,$$

so that the elimination of λ', μ', ν', ϱ' will lead to the same results as before.

In general it may be observed that each combination of s, t, v which is possible without logical contradiction, gives, on substituting for s, t, v... their expressions in the simple terms of the original problem, either a single combination of those simple terms, or a sum of such combinations; but the same combination of those simple terms will not arise from two different combinations of s, t... It is clear from this that the systems of united equations and inequations arising in the two forms of the problem will be related in the following manner, viz.—

For each positive quantity λ' in the one set, there will exist either a single positive quantity λ, or a sum of such quantities $\lambda_1+\lambda_2+$&c. in the other; but each such sum is inseparable, and the elements it is composed of are distinct from those of any other sum arising from any other of the quantities λ'... It is evident, then, that the final results of elimination will be the same. The same formal processes which eliminate single quan-

* Laws of Thought, p. 267.

tities in the one case, will eliminate the corresponding single quantities, or sums of single quantities, in the other.

Simplification of the General Equations for the Solution of Questions in the Theory of Probabilities.

Let us express the system (I.) in the form

$$\frac{V_s}{V}=p,\quad \frac{V_t}{V}=q,\ \&c.,$$

$$u=\frac{A+cC}{V},$$

and let us suppose the quantities p, q... (and therefore s, t...) to be n in number. Then all the terms in V will be composed of products of s, t... $\bar{s}$, $\bar{t}$..., each term involving either s or $\bar{s}$, either t or $\bar{t}$, &c., but not the combinations $s\bar{s}$, $t\bar{t}$, &c. Each term is therefore homogeneous and of the nth degree.

It follows, therefore, that if we divide the numerator and denominator of each of the first members of the above system by $\bar{s}\,\bar{t}\,\bar{v}$..., and then make

$$\frac{s}{\bar{s}}=x_1,\quad \frac{t}{\bar{t}}=x_2,\quad \frac{v}{\bar{v}}=x_3,\ \&c.,$$

and if at the same time we, for symmetry, change p, q, r... into p_1, p_2, ..p_n, the system will assume the following form,—

$$\frac{V_1}{V}=p_1,\quad \frac{V_2}{V}=p_2\cdots\frac{V_n}{V}=p_n,$$

$$u=\frac{A+cC}{V},$$

in which V, A, C are formed from their former values by suppressing $\bar{s}$, $\bar{t}$, $\bar{v}$, &c., or, which is the same thing, changing each of them into unity, and then changing s, t, v... into x_1, x_2, x_3..., while V_1 consists of those terms of V which contain x_1, V_2 of those which contain x_2, and so on.

In its new form V is a rational and entire function of x_1, x_2, ...x_n not involving powers of those quantities, and with all its coefficients equal to unity. Again, as s, t, &c. are from the theory of their origin required to be positive proper fractions, x_1, x_2, ... x_n are, from the nature of their connexion with s, t..., required to be positive quantities. And it is sufficient that x_1, x_2, ... x_n be determinable as positive quantities in order that s, t... may be determinable as positive fractions.

Now we shall proceed to show that x_1, x_2, ... x_n are determinable as positive quantities precisely when p_1, p_2, ...p_n satisfy the conditions of possible experience. We shall further show, as a consequence of this, that the value of the probability sought, when determined by the General Rule, will, under the same conditions, lie within such limits as if it were itself given by the same experience. In the order of this proof, we shall first demonstrate the theorems of pure Analysis upon which the conclusions depend, then in a distinct section make the particular application.

Analytical Theorems relating to Functional Determinants and Systems of Algebraic Equations.

A symmetrical determinant may be conveniently expressed in the form

$$\begin{vmatrix} A_1 & A_{12} & \dots & A_{1n} \\ A_{21} & A_2 & \dots & A_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ A_{n1} & A_{n2} & \dots & A_n \end{vmatrix} \quad \dots\dots\dots\dots \quad (\text{I.})$$

the conditions of symmetry being

$$A_{ij}=A_{ji},\ A_{ii}=A_i.$$

It is desirable to employ fixed language in referring to this. We shall therefore call the quantities $A_1, A_2 \dots A_n$ the 'principal elements,' and the diagonal series of terms which they form the 'principal diagonal.' The elements A_{ij}, when i and j differ, we shall call 'subordinate elements.' The element A_i, together with all the subordinate elements which occur upon the same horizontal or vertical line of the determinant, we shall designate the 'i-system of elements.' Lastly, in comparing two rows or two columns of elements together, those elements will be said to correspond which occupy the same numerical place in their respective rows or columns.

The following Lemma will next be established.

Lemma.—A symmetrical determinant expressed in the form (I.) will be unaltered in value, if from each subordinate element of its i-system we subtract the corresponding element of its j-system multiplied by a quantity λ, which is invariable for the same system,—and for the principal element A_i substitute $A_i-2\lambda A_{ij}+\lambda^2 A_j$.

It is known that a determinant vanishes if two of its lines or columns are identical, and it is known as a consequence of this that if from a particular line or column of a determinant the corresponding elements of another line or column, multiplied each by the same constant, are subtracted, the determinant is unaltered in value. From the ith line of the above symmetrical determinant subtract, term by term, λ times the jth line, and then from the ith column of the resulting determinant subtract λ times the jth column. As respects any subordinate element, the result will obviously accord with the statement in the Lemma. But the element A_i will be successively converted into

$$A_i-\lambda A_{ji}$$

$$(A_i-\lambda A_{ji})-\lambda(A_{ij}-\lambda A_j).$$

The last expression, since $A_{ji}=A_{ij}$, is reducible to

$$A_i-2\lambda A_{ij}+\lambda^2 A_j.$$

Upon this property the demonstration of the following general proposition will be founded.

PROPOSITION I.

Let the symmetrical determinant (I.) *possess the following properties, viz.:—*

1st. *That all its elements are linear homogeneous rational functions of certain quantities a, b, c, &c., unlimited in number.*

2ndly. *That if the coefficients of any one of these quantities a in the elements of any particular line or column taken in order are $\alpha_1, \alpha_2, \ldots \alpha_n$, and in any other line or column $\beta_1, \beta_2, \ldots \beta_n$, then these two series of quantities are respectively proportional.*

3rdly. *That the principal terms $A_1, A_2, \ldots A_n$ are positive*, i. e. *that the coefficients of all the quantities a, b, c, &c. which appear in these terms are positive.*

Then the developed determinant will be itself positive, and will consist of products of the quantities a, b, c, &c. without powers, each product affected by a positive sign.

First, it may be observed that any letter a of the set $a, b, c \ldots$ which appears in the subordinate term A_{ij} will appear in both the principal terms A_i, A_j.

For let m be the coefficient of a in A_{ij}, and therefore also in A_{ji}; let l be the coefficient of a in A_i, and n its coefficient in A_j. Thus to the elements A_i, A_{ji} in the i-column correspond A_{ji}, A_j in the j-column. Hence, by the definition of the determinant,

$$l : m :: m : n,$$
$$\therefore \quad m^2 = ln,$$

which implies that neither l nor n vanishes, so that a appears in A_i and A_j.

Secondly, we shall show that the determinant can, without alteration of its final developed value, be reduced to a form in which any letter a of the system $a, b, c \ldots$ shall appear in only one system of elements, and therefore only in the principal term of that system, since every subordinate term is common to *two* systems.

Let us suppose a to be contained in two at least of the systems of elements, and for convenience of expression, let these be the 1-system and the n-system. Let, then, $\alpha_1, \alpha_2, \ldots \alpha_n$ be the successive coefficients of a in $A_1, A_{21}, \ldots A_{n1}$, and therefore, by definition of the determinant, $\lambda\alpha_1, \lambda\alpha_2, \ldots \lambda\alpha_n$, its coefficients in $A_{n1}, A_{n2}, \ldots A_n$. Any of the quantities $\alpha_1, \alpha_2, \ldots \alpha_n$ may be 0. But by the Lemma above demonstrated the determinant may, without alteration of value, be reduced to the following form, viz.:—

$$\begin{vmatrix} A_1, & A_{12}, \ldots & A_{1n} - \lambda A_1 \\ A_{21}, & A_2, \ldots & A_{2n} - \lambda A_{21} \\ \cdot & \cdot & \cdot \\ A_{n1} - \lambda A_1, & A_{n2} - \lambda A_{12} \ldots & A_n - 2\lambda A_{n1} + \lambda^2 A_1 \end{vmatrix} \quad \ldots\ldots\ldots \quad (B.)$$

Now in the determinant thus transformed the quantity a will no longer occur in the n-system.

This is obvious with respect to the subordinate elements of that system. With respect to the principal element, we observe that the coefficient of a is

in A_1, equal to α_1,
in A_{n1}, equal to $\lambda\alpha_1$,
. . . . . .
in A_n, equal to $\lambda \times \lambda\alpha_1$ or $\lambda^2\alpha_1$,

whence the coefficient of a in $A_n - 2\lambda_{n1} + \lambda^2 A_1$ is equal to 0.

Thus a has been eliminated from the n-system, and as the process has not affected any elements but those which belong to the n-system, it will not affect the relations under which a enters into the other systems.

Consider then any other quantity b in the set a, b, c, then by hypothesis the coefficients of b in any line or column of elements

$$A_{i1}, A_{i2}, .. A_{in}, \text{ or } A_{1i}, A_{2i}, .. A_{ni}$$

may be represented by

$$\mu_i\beta_1, \ \mu_i\beta_2, .. \mu_i\beta_n,$$

$\beta_1, \beta_2, .. \beta_n$ being an arbitrary set of quantities which are the same for all lines or columns, while μ_i differs for different lines or columns, and vanishes for those in which b does not enter.

It is to be noted that as $A_{ij}=A_{ji}$, we have in general

$$\mu_i\beta_j=\mu_j\beta_i,$$

while as the principal elements of the determinant (I.) are positive, we have always $\mu_i\beta_i=$a positive quantity.

Now reverting to the derived determinant (B.), we see that its ith line or column of elements will be

$$A_{i1}, \ A_{i2}, \ldots \ A_{in}-\lambda A_{i1},$$

and its jth line or column

$$A_{j1}, \ A_{j2}, \ldots \ A_{jn}-\lambda A_{j1},$$

supposing i and j to be both less than n.

In these lines or columns the successive coefficients of b will therefore be

$$\mu_i\beta_1, \ \mu_i\beta_2, \ \ldots \ \mu_i\beta_n-\lambda\mu_i\beta_1,$$
$$\mu_j\beta_1, \ \mu_j\beta_2, \ \ldots \ \mu_j\beta_n-\lambda\mu_j\beta_1,$$

which stand to each other in the constant ratio $\mu_i : \mu_j$.

Now let $j=n$. The coefficients of b in the nth line or column of (B.) are obviously

$$\mu_n\beta_1-\lambda\mu_1\beta_1, \ \mu_n\beta_2-\lambda\mu_1\beta_2, \ \ldots \ \mu_n\beta_n-2\lambda\mu_1\beta_n+\lambda^2\mu_1\beta_1,$$

of which the last term may be reduced as follows,

$$\mu_n\beta_n-2\lambda\mu_1\beta_n+\lambda^2\mu_1\beta_1=\mu_n\beta_n-\lambda\mu_1\beta_n-\lambda\mu_n\beta_1+\lambda^2\mu_1\beta_1=(\mu_n-\lambda\mu_1)(\beta_n-\lambda\beta_1);$$

so that the series of coefficients of b becomes

$$(\mu_n-\lambda\mu_1)\beta_1, \ (\mu_n-\lambda\mu_1)\beta_2 \ \ldots \ (\mu_n-\lambda\mu_1)(\beta_n-\lambda\beta_1),$$

and they are now seen to stand to those of b in the i-line and column in the constant ratio $\mu_n-\lambda\mu_1 : \mu_i$.

We have, lastly, to prove that the new principal element $A_n-2\lambda A_{1n}+\lambda^2 A_1$ is positive.

Let N be the coefficient of any one of the quantities $a, b, c \ldots$ in the above element, L its coefficient in the principal element A_i, and M its coefficient in each of the subordinate elements common to the two systems of which the above are the respective

principal elements, viz. in $A_{in}-\lambda A_{i1}$ and $A_{ni}-\lambda A_{1i}$. Then, by what has already been proved,

$$L : M :: M : N,$$

$$\therefore \quad M^2 = LN;$$

but L is positive; therefore N is so, and the principal element in question consists wholly of positive terms.

The above demonstration shows that the elimination of a from the n-system produces a new determinant equivalent to the original one, and in which the characters noted in the original one still remain. Should a occur in any other system or systems of elements of the new determinant beside the 1-system, it can, by repetitions of the same process, be eliminated thence. Ultimately, then, it will only remain in the 1-system, and therefore only in the principal term of that system. Again, as it enters that term in the first degree, it follows that the developed determinant will involve only the first power of a. Hence, as a may represent any of the quantities $a, b, c, \ldots$, it is seen that no powers, but only products of these quantities, will appear in the developed determinant.

Let us represent the determinant, after the elimination of a from all the elements but A_1, in the form

$$\begin{vmatrix} A_1 & B_{12} \ldots & B_{1n} \\ B_{21} & C_2 \ldots & C_{2n} \\ \cdot & \cdot \quad \cdot & \cdot \quad \cdot \\ B_{n1} & C_{n2} \ldots & C_n. \end{vmatrix}$$

Now let ah_1 represent that term in A_1 which involves a. Then the portion of the determinant which involves a will be

$$ah_1 \begin{vmatrix} C_2 \ldots & C_{2n} \\ \cdot \quad \cdot & \cdot \quad \cdot \\ C_{n2} \ldots & C_n. \end{vmatrix}$$

And here it is to be observed that ah_1 is positive, while the new determinant to which it is attached as a coefficient possesses all the characters of the old one. This determinant we can therefore transform in the same way, so as to eliminate any other letter b from all but a single principal element, which we shall suppose to contain it in a term bh_2. That portion of the original determinant which involves ab will therefore assume the form

$$abh_1h_2 \begin{vmatrix} D_3 \ldots & D_{3n} \\ \cdot \quad \cdot & \cdot \quad \cdot \\ D_{n3} \ldots & D_n. \end{vmatrix}$$

Ultimately, then, as the result of such processes continued, the portion of the original determinant which involves any particular combination of n letters selected from $a, b, c \ldots$ will consist of the product of a series of positive terms, each of which has appeared in some residual principal element. Every such combination being positive, it follows that the determinant itself consists solely of positive terms.

PROPOSITION II.

If V *be any rational entire function of the n variables* $x_1, x_2, \ldots x_n$, *but involving no powers of those variables above the first, and if, further, all the different terms of* V *have positive signs, then the determinant*

$$\begin{vmatrix} V & V_1 & V_2 & \ldots & V_n \\ V_1 & V_1 & V_{12} & \ldots & V_{1n} \\ V_2 & V_{21} & V_2 & \ldots & V_{2n} \\ . & . & . & . & . \\ V_n & V_{n1} & V_{n2} & \ldots & V_n \end{vmatrix}$$

in which V_i *denotes the sum of the terms in* V *which contain* x_i, *and* V_{ij} *the sum of the terms in* V *which contain* x_i, x_j, *will, when developed as a rational and entire function of* $x_1, x_2, \ldots x_n$, *consist wholly of terms with positive coefficients.*

From the definition it is plain that in general

$$V_{ij}=V_{ji}, \quad V_{ii}=V_i,$$

whence the above determinant is symmetrical.

Again, all its elements are homogeneous linear functions of the terms in V.

Again, if $\alpha, \alpha_1, \alpha_2, \ldots \alpha_n$ represent the successive coefficients of any one of the terms of V in any row or column of the determinant, and $\beta, \beta_1, \beta_2 \ldots \beta_n$ the successive corresponding coefficients of the same term in any other row or column of the determinant, the one series of coefficients shall be proportional to the other.

Let us compare the first column and the i-column headed with the element V_i. Selecting any term in V, suppose it to contain x_i, then in whatever element of the first column that term is found, it will be found in a corresponding element of the i-column, and in each case with unity for its coefficient, since all the elements are mere collections of terms from V. But when it is not found in a particular element of the first column, it will not be found in the corresponding element of the i-column. The entire series of coefficients in the one being then the same as that in the other, the common ratio of the corresponding terms is unity.

Suppose, secondly, that the proposed term is found in V and not in V_i; then in all the elements of the i-column its coefficient is 0, so that the series of coefficients in the i-column might be formed from those in the first column by multiplying the latter successively by 0. This again represents a common ratio.

The same reasoning may be applied to the comparison of any two columns of the determinant. Thus in comparing the i-column and the j-column:—terms of V which contain both x_i and x_j will be found in corresponding elements of both columns—terms which contain x_i but not x_j will be wholly absent from the j-column. Thus in all cases if $\alpha, \alpha_1, \alpha_2, \ldots \alpha_n$ represent the coefficients of a term of V in one column, its coefficients in any other column, taken in the same order, will be of the form $\lambda\alpha, \lambda\alpha_1, \lambda\alpha_2 \ldots \lambda\alpha_n$, the coefficient λ being either 1 or 0.

Lastly, the principal elements consist, as do all the elements, of positive terms.

Therefore by the last proposition the developed determinant will consist of products (without powers higher than the first) of different terms of V, and the coefficients of all such products will be positive.

Therefore the determinant will be expressible as a rational entire function of $x_1, x_2, \ldots x_n$ with positive coefficients.

The rapidity with which the complexity of the determinant increases as the number of variables increases is remarkable. For example, if $n=2$ and $V=axy+bx+cy+d$, the determinant is

$$\begin{vmatrix} axy+bx+cy+d & axy+bx & axy+cy \\ axy+bx & axy+bx & axy \\ axy+cy & axy & axy+cy; \end{vmatrix}$$

and its calculated value will be found to be

$$abcx^2y^2+abdx^2y+acdxy^2+bcdxy,$$

consisting of four positive terms.

But if $n=3$ and

$$V=axyz+byz+cxz+dxy+ex+fy+gz+h,$$

the developed determinant will consist of fifty-eight positive terms. Its calculated value will be found in the Memoir on Testimonies and Judgments.

Proposition III.

The functions V, V_1, $V_2, \ldots V_n$ *being defined as above, if* V *be complete in form,* i. e. *if it consist of all the terms which according to definition it can contain, each with a positive coefficient, then the system of equations*

$$\frac{V_1}{V}=p_1, \quad \frac{V_2}{V}=p_2 \ldots \frac{V_n}{V}=p_n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1.)$$

will, when $p_1, p_2, \ldots p_n$ *are proper fractions, admit of one solution, and only one, in positive values of* $x_1, x_2, \ldots x_n$.

We shall show, first, that the above proposition is true when $n=1$, secondly, that on the hypothesis that it is true for $n-1$ variables, it is true for n variables. Hence it will follow that it is true generally.

Suppose $n=1$. Then $V=ax_1+b$, whence the system (1.) reduces to the single equation

$$\frac{ax_1}{ax_1+b}=p_1,$$

$$\therefore \qquad x_1=\frac{bp_1}{a(1-p_1)},$$

whence, since a and b are positive, and p is a positive fraction, x_1 is positive.

Thus the proposition is true when $n=1$.

Now, let $x_1=0$, and let $x_2, x_3 \ldots x_n$ be determined to satisfy the last $n-1$ equations of the system (1.). These $n-1$ equations will, when $x_1=0$, form a system of the same

nature with respect to the $n-1$ variables $x_2, x_3 \ldots x_n$, as (1.) is with respect to the n variables $x_1, x_2, \ldots x_n$. This will be at once seen by taking any particular example. Hence by hypothesis $x_2, x_3 \ldots x_n$ will be determinable as positive quantities, and their values substituted in the first member of the first equation of (1.) will reduce it to the form

$$\frac{Ax_1}{Ax_1+B},$$

A and B being finite and positive. Hence the function $\frac{V_1}{V}$ will become 0.

Secondly, let any finite positive value be assigned to x_1. The last $n-1$ equations of the system (1.) will again form a system of the same nature as before, and will by hypothesis determine a set of finite positive values for $x_2, x_3, \ldots x_n$. These values again substituted in $\frac{V_1}{V}$, will give to it again the form

$$\frac{Ax_1}{Ax_1+B},$$

A and B being finite and positive. Hence as x_1 is finite and positive, $\frac{V_1}{V}$ will be a positive fraction.

Lastly, let x_1 be infinite. Still the last $n-1$ equations of the system (1.) will assume the same form as before. Determining thence $x_2, x_3 \ldots x_n$, and substituting in $\frac{V_1}{V}$, we have

$$\frac{V_1}{V}=\frac{Ax_1}{Ax_1+B},$$

in which A and B are finite and positive and x_1 is infinite. Hence $\frac{V_1}{V}=1$. It is seen then that as x_1 varies from 0 to infinity, $x_2, x_3, \ldots x_n$ being at the same time always by hypothesis determined to satisfy the last $n-1$ equations of the system (1.), the function $\frac{V_1}{V}$ will vary from 0 through positive fractional values to unity. It is manifest, too, that it varies continuously. If then it vary by continuous *increase*, it will once, and only once in its change, become equal to p_1, and the whole system of equations thus be satisfied together. I shall show that it does vary by continuous *increase*.

If it vary continuously from 0 to 1 and not by continuous increase, it must in the course of its variation assume at least once a maximum or minimum value. Let us then seek the condition of possibility of

$$\frac{V_1}{V} = \text{a maximum or minimum},$$

the variables being subject to the relations

$$\frac{V_2}{V}=p_2, \quad \frac{V_3}{V}=p_3 \ldots \frac{V_n}{V}=p_n.$$

Here, proceeding in the usual way by differentiation, we have

$$\frac{VdV_1-V_1dV}{V^2}=0, \quad \frac{VdV_2-V_2dV}{V^2}=0, \ldots \frac{VdV_n-V_ndV}{V^2}=0,$$

or

$$\frac{dV}{V}=\frac{dV_1}{V_1}=\frac{dV_2}{V_2}\cdots=\frac{dV_n}{V_n}.$$

Let the common value of these fractions be represented by $-dt$, then we have a system of $n+1$ equations of which the first is

$$Vdt+dV=0,$$

while the n others are of the type

$$V_i dt+dV_i=0.$$

The complete system, therefore, on effecting the total differentiations, becomes

$$Vdt+\frac{dV}{dx_1}dx_1\ldots+\frac{dV}{dx_n}dx_n=0,$$

$$V_1 dt+\frac{dV_1}{dx_1}dx_1\ldots+\frac{dV_1}{dx_n}dx_n=0,$$

$$\cdots\cdots\cdots\cdots$$

$$V_n dt+\frac{dV_n}{dx_1}dx_1\ldots+\frac{dV_n}{dx_n}dx_n=0.$$

Now from the nature of the function V we have

$$\frac{dV}{dx_i}=\frac{V_i}{x_i},\quad \frac{dV_i}{dx_i}=\frac{V_i}{x_i},\quad \frac{dV_i}{dx_j}=\frac{V_{ij}}{x_j},$$

so that the above equations become

$$Vdt+V_1\frac{dx_1}{x_1}+V_2\frac{dx_2}{x_2}\ldots+V_n\frac{dx_n}{x_n}=0,$$

$$V_1 dt+V_1\frac{dx_1}{x_1}+V_{12}\frac{dx_2}{x_2}\ldots+V_{1n}\frac{dx_n}{x_n}=0,$$

$$V_2 dt+V_{21}\frac{dx_1}{x_1}+V_2\frac{dx_2}{x_2}\ldots+V_{2n}\frac{dx_n}{x_n}=0,$$

$$\cdots\cdots\cdots\cdots$$

$$V_n dt+V_{n1}\frac{dx_1}{x_1}+V_{n2}\frac{dx_2}{x_2}\ldots+V_n\frac{dx_n}{x_n}=0,$$

and the elimination of dt, $\frac{dx_1}{x_1}$, $\frac{dx_2}{x_2}\ldots\frac{dx_n}{x_n}$ from these equations gives the sought condition of possibility of a maximum value of $\frac{V_1}{V}$, consistently with the satisfaction of the last $n-1$ equations of the system (1.).

This condition is therefore expressed by the equation

$$\begin{vmatrix} V & V_1 & V_2 \ldots V_n \\ V_1 & V_1 & V_{12}\ldots V_{1n} \\ V_2 & V_{21} & V_2 \ldots V_{2n} \\ \ldots & \ldots & \ldots \quad \ldots \\ V_n & V_{n1} & V_{n2}\ldots V_n \end{vmatrix}=0.$$

But we have already seen (Prop. II.) that the first member of this equation is essentially positive for positive values of $x_1, x_2 \ldots x_n$. Hence the function $\frac{V_1}{V}$ varies by continuous increase, and on the hypothesis that the proposition to be proved is true for $n-1$ variables, it is true for n variables.

Therefore, connecting this with the former result, the proposition is true universally.

Proposition IV.

If V *be an incomplete function, some of the terms belonging to the complete form being wanting, but the terms present having their coefficients positive, it will in general be necessary not only that the quantities $p_1, p_2, \ldots p_n$ should be positive fractions, but also that they should satisfy certain inequations of the form*

$$a_1p_1+a_2p_2\ldots+a_np_n+b \geqq 0,$$

in order that the system

$$\frac{V_1}{V}=p_1, \quad \frac{V_2}{V}=p_2 \ldots \frac{V_n}{V}=p_n \quad \ldots\ldots\ldots \quad (1.)$$

may admit of a solution in positive values of $x_1, x_2 \ldots x_n$.

For let $Ax_r x_s x_t \ldots$ be any term in V, A being a constant which is positive in all the terms, but which may be different in the different terms. Suppose that in V_i there exist e terms like the above, and let the several ratios of these terms to V be denoted by $\lambda_1, \lambda_2 \ldots \lambda_e$. Then the ith equation of the system (1.) will become

$$\lambda_1+\lambda_2\ldots+\lambda_e=p_i, \quad \ldots\ldots\ldots \quad (2.)$$

and the system (1.) will be converted into a system of n equations of this nature. We will suppose that there exist m distinct quantities of the nature of $\lambda_1, \lambda_2 \ldots \lambda_e$ in the first members of this transformed system, and we will represent these by $\lambda_1, \lambda_2 \ldots \lambda_m$. Then, if these constitute all the ratios of the separate terms of V to V itself, we have a new equation,

$$\lambda_1+\lambda_2\ldots+\lambda_m=1. \quad \ldots\ldots\ldots \quad (3.)$$

If they do not constitute all those separate ratios, we have, on the contrary, an inequation,

$$\lambda_1+\lambda_2\ldots+\lambda_m \leqq 1. \quad \ldots\ldots\ldots \quad (4.)$$

Lastly, the condition that $\lambda_1, \lambda_2 \ldots \lambda_m$ are positive fractions, gives the inequations

$$\lambda_1 \geqq 0, \ \lambda_2 \geqq 0 \ldots \lambda_m \geqq 0. \quad \ldots\ldots\ldots \quad (5.)$$

The conditions $\lambda_1 \leqq 1$, &c. are already implied in (3.) or (4.).

The λ quantities are thus subject to a system of united *equations* and *inequations*, from which they must be eliminated by the method already explained.

The result of such elimination will be a final system of inequations connecting $p_1, p_2, \ldots p_n$. Equations connecting these quantities can only present themselves when the equations of the original system are not independent, or, which really falls under

the same hypothesis, when one or more of the variables $x_1, x_2 \ldots x_n$ is wholly absent from that system. Thus if x_1 were a common factor of all the terms of V, it would divide out from the numerators and denominators of the system, which would thus become a system of n simultaneous equations connecting the $n-1$ variables $x_2, x_3 \ldots x_n$. Considered with reference to these variables, therefore, the equations of the system would not be independent.

All resulting inequations will be capable of expression under the one general form,

$$a_1p_1+a_2p_2 \ldots +a_np_n+b \overline{\overline{>}} 0,$$

the coefficients $a_1, a_2, \ldots a_n$ and b being positive, negative, or vanishing, numerical constants. For any inequation which presents itself in the form

$$a'_1p_1+a'_2p_2 \ldots +a'_np_n+b \overline{\overline{<}} 1$$

may be transformed into

$$-a'_1p_1 \ldots -a'_np_n+1-b \overline{\overline{>}} 0.$$

Again, the general inequation

$$a_1p_1+a_2p_2 \ldots +a_np_n+b \overline{\overline{>}} 0$$

determines an inferior limit of p_1 when a_1 is positive, and a superior limit of p_1 when a_1 is negative.

For in the former case we have

$$p_1 \overline{\overline{>}} -\left(\frac{a_2}{a_1}p_1 \ldots +\frac{a_n}{a_1}p_n+\frac{b}{a_1}\right),$$

the second member of which is an inferior limit of p_1; and it will be observed that the calculated value of this member may be positive, as there is no general restriction on the signs of $a_2, \ldots a_n, b$.

In the latter case, changing a_1 into $-a'_1$, and observing that a_1 is positive, we have

$$p_1 \overline{\overline{<}} \frac{a_2}{a'_1}p_2+\frac{a_3}{a'_1}p_3 \ldots +\frac{a_n}{a'_1}p_n+\frac{b}{a'_1},$$

the second member of which is a superior limit of p_1.

Lastly, the final system of inequations is totally independent of the numerical value of the coefficients of V. The only restriction is that these coefficients are positive.

Proposition V.

Let V *be incomplete in form; then, provided that the equations*

$$\frac{V_1}{V}=p_1,\ \frac{V_2}{V}=p_2 \ldots \frac{V_n}{V}=p_n \quad \ldots\ldots\ldots\ldots \quad (1.)$$

are independent with respect to the quantities $x_1, x_2, \ldots x_n$, *and that the inequations of condition deducible by the last proposition are satisfied, the equations will admit of one solution, and only one, in positive finite values of* $x_1, x_2, \ldots x_n$.

The proof of this proposition will, in its general character, resemble the proof of

Proposition III. It will be shown that when we assign to x any value *between* the limits 0 and infinity, the quantities $x_2, x_3 \ldots x_n$ will admit of determination from the last $n-1$ equations of the system as positive finite quantities, and the function $\frac{V_1}{V}$ will receive a value falling within the limits assigned by Proposition IV. to the quantity p_1; that when x_1 is equal to 0 or infinity, $x_2, x_3 \ldots x_n$ will admit of determination either as positive finite quantities, or as limits (0 and ∞) of such quantities; and that these values together will give to $\frac{V_1}{V}$ a value coinciding with the highest of the inferior, or lowest of the superior limits of p_1, as determined by Proposition IV.; that when x_1 varies from 0 to ∞, $x_2, x_3 \ldots x_n$ being determined as above, $\frac{V_1}{V}$ will vary by continuous increase from the highest of the inferior to the lowest of the superior limits of p_1, and once in its variation become equal to p_1. Thus the truth of the proposition for n variables will flow necessarily from its assumed truth for $n-1$ variables. And on this ground it will be shown that it may ultimately be reduced to a direct dependence upon Proposition III.

In the system (1.) let x_1 receive any finite positive value, and let V by the substitution of this value become U; the last $n-1$ equations of (1.) will thus assume the form

$$\frac{U_2}{U}=p_2,\ \frac{U_3}{U}=p_3 \ldots \frac{U_n}{U}=p_n, \quad \ldots\ldots\ldots \quad (2.)$$

in which the quantities $p_2, p_3 \ldots p_n$ satisfy the conditions to which the direct application of Proposition IV. to this reduced system of equations would lead.

For what is important to notice in the change from V to U is, that any two terms in V which differ only in that one contains x_1 and the other does not, reduce to a single term in U. The effect of the change upon the primary system of equations and inequations formed in the application of Proposition IV. to the system (1.) is the following:—

1st. The equation between $\lambda_1, \lambda_2 \ldots$ derived from the first equation of (1.) will be annulled.

2ndly. In the remaining equations connecting $\lambda_1, \lambda_2 \ldots$ some *pairs* of those quantities may be replaced by single quantities, with corresponding changes in the inequations. Thus if $\lambda_1+\lambda_2$ be replaced by μ, the inequations

$$\lambda_1 \overset{=}{>} 0,\ \lambda_2 \overset{=}{>} 0$$

will be replaced by what they before *implied*, viz.

$$\mu \overset{=}{>} 0.$$

But these changes do not affect the truth of the relations, or introduce any new relations. They cannot, therefore, lead to any new final conditions. The conditions connecting $p_2, p_3 \ldots p_n$, in accordance with Proposition IV. in the system (2.), must have been already involved in the equations connecting $p_1, p_2 \ldots p_n$ in the system (1.).

Hence by hypothesis the system (2.) gives one set of positive finite values of

$x_2, x_3 \ldots x_n$ corresponding to the assumed positive finite value of x_1. And these values together make $\frac{V_1}{V}$ a positive proper fraction. We may notice that, representing $\frac{V_1}{V}$ under the form

$$\frac{Ax_1}{Ax_1+B},$$

it cannot be that either A or B is wanting so as to reduce $\frac{V_1}{V}$ to the value 0 or 1. For if A were wanting, V would not contain x_1 at all, as by hypothesis it does; and if B were wanting, V would contain x_1 in every term. Thus x_1 would divide out from the system (1.), which would thus become a system of n equations between $n-1$ variables, and would cease to be independent, as by hypothesis it is.

But when $x_1=0$, or $x_1=$ infinity, the form of V, considered as a function of $x_2, x_3 \ldots x_n$, will not generally be the same as in the case last considered; and the conditions connecting $p_2, p_3, \ldots p_n$ will no longer be such that we can affirm the possibility of deducing from the last $n-1$ equations of the system (1.), as transformed, positive finite values of $x_2, x_3, \ldots x_n$.

The theory of this case depends upon a remarkable transformation.

The most general form of the inequations of condition connecting $p_1, p_2, \ldots p_n$, as determined by Proposition IV., is

$$a_1p_1+a_2p_2 \ldots +a_np_n+b \overline{\gtrless} 0. \quad \ldots \ldots \ldots \quad (3.)$$

Hence, from the nature of the system (1.), it follows that the function

$$a_1V_1+a_2V_2 \ldots +a_nV_n+bV \quad \ldots \ldots \ldots \quad (4.)$$

must consist wholly of positive terms. Therefore V must consist of terms which would either appear in the development of the above function with positive signs, or not appear in it at all. Let $Ax_rx_sx_t \ldots$ be any term of V. Then, as the coefficient of this term in (4.) would be

$$a_rA+a_sA+a_tA \ldots +bA,$$

and as A is positive, we have

$$a_r+a_s+a_t \ldots +b \overline{\gtrless} 0,$$

a general condition which determines not what terms have actually entered, but what could alone possibly have entered into the constitution of V.

From the system (1.) we have

$$\frac{a_1V_1+a_2V_2 \ldots +a_nV_n+bV}{V}=a_1p_1+a_2p_2 \ldots +a_np_n+b.$$

Hence if we write

$$a_1V_1+a_2V_2 \ldots +a_nV_n+bV=H,$$

we have

$$\frac{H}{V}=a_1p_1+a_2p_2 \ldots +a_np_n+b, \quad \ldots \ldots \ldots \quad (5.)$$

an equation by which we may replace any one of the equations of the system (1.), and

which has the peculiarity that for every term $Ax_r\,x_s\,x_t\ldots$ which appears in the numerator H the particular condition

$$a_r+a_s+a_t\ldots+b>0$$

is satisfied.

Let K be the aggregate of those terms in V for which the remaining particular condition

$$a_r+a_s+a_t\ \ldots+b=0$$

is satisfied. Then $V=H+K$. If we now substitute (5.) in place of the first equation of the system (1.) and then write $H+K$ for V, H_1+K_1 for V_1, &c., the system becomes converted into the following one, viz.

$$\frac{H}{H+K}=a_1p_1+a_2p_2\ldots+a_np_n+b,\ \frac{H_2+K_2}{H+K}=p_2,\ \frac{H_3+K_3}{H+K}=p_3,\ldots\frac{H_n+K^n}{H+K}=p_n.\quad .\quad . \quad (6.)$$

Now let us transform the above equations by assuming

$$x_2=x_1^{\frac{a_2}{a_1}}y_2,\ x_3=x_1^{\frac{a_3}{a_1}}y_3\ldots x_n=x_1^{\frac{a_n}{a_1}}y_n.$$

The general type of these equations is

$$x_i=x_1^{\frac{a_i}{a_1}}y_i,$$

and it includes the particular case of $i=1$, provided that we suppose, as we shall do, $y_1=1$.

Then representing, as before, any term of V by $Ax_r\,x_s\,x_t\ldots$, we have

$$Ax_r\,x_s\,x_t\ldots=Ax_1^{\frac{a_r+a_s+a_t\ldots}{a_1}}y_r\,y_s\,y_t\ldots$$

Let this substitution be made in the different terms both of the numerators and denominators of the fractions which form the first members of the above system, and then let each numerator and denominator be multiplied by $x_1^{\frac{b}{a_1}}$. The result will be the same as if for each term $Ax_r\,x_s\,x_t\ldots$ in numerator or denominator we substituted the term

$$Ax_1^{\frac{a_r+a_s+a_t\ldots+b}{a_1}}y_r\,y_s\,y_t\ldots$$

In considering the effect of this transformation we will first suppose a_1 positive, and afterwards suppose it negative.

Case 1; the coefficient a_1 positive. Here, since for all the terms in H and in H_2, $H_3\ldots H_n$ we have

$$\frac{a_r+a\ +a_t\ldots+b}{a_1}>0,$$

all such terms in the transformed equations will be affected with positive powers of x_1.

And since for all terms in K, $K_2,\ldots K_n$ we have

$$\frac{a_r+a_s+a_t\ldots+b}{a_1}=0,$$

all such terms in the transformed equations will be free from x_1.

Now let $$a_1p_1+a_2p_2\ldots+a_np_n+b=0.$$

This, as a_1 is positive, is to suppose that p_1 coincides with one of its own inferior limits. We must suppose this to be the highest of those limits, as otherwise some of the other limiting conditions would be violated. Now, since all the terms in H are affected with positive powers of x_1, while those in K do not contain x_1, the first equation of the system (6.) will be satisfied by $x_1=0$, provided that the remaining $n-1$ equations give finite positive values for $y_2\ldots y_n$. But the vanishing of x_1 reduces these equations to the form

$$\frac{K_2}{K}=p_2,\ \frac{K_3}{K}=p_3,\ldots\ \frac{K_n}{K}=p_n. \quad \ldots\ldots\ldots \quad (7.)$$

It is therefore necessary to show that $p_2, p_3\ldots p_n$ in this system are actually subject to the conditions to which the application of the method of Proposition IV. to the system itself would lead.

The n quantities $p_1, p_2\ldots p_n$ are by hypothesis subject to the conditions furnished by the application of the method of Proposition IV. to the original system (1.). In applying this method each of the original equations yields an equation of the form

$$\lambda_1+\lambda_2\ldots+\lambda_e=p_i; \quad \ldots\ldots\ldots\ldots \quad (8.)$$

and to the equations thus formed are added the inequations

$$\lambda_1+\lambda_2\ldots+\lambda_m\overline{\overline{<}}1,$$
$$\lambda_1\overline{\overline{>}}0,\ \lambda_2\overline{\overline{>}}0,\ \ldots\lambda_m\overline{\overline{>}}0;$$

$\lambda_1, \lambda_2\ldots\lambda_m$ having reference to the whole system of original equations.

Now the satisfaction of the equation

$$\frac{H}{H+K}=0$$

by the value $x_1=0$, involves the vanishing of all those quantities of the system $\lambda_1, \lambda_2\ldots\lambda_m$, which are derived from terms in V that are also found in H. Hence the λ quantities that do not vanish are those derived from terms in V which appear in K.

Again, the condition

$$a_1p_1+a_2p_2\ldots+a_np_n+b=0$$

shows that the system of equations of which (8.) is the type are not independent. They must, under the particular circumstances of the case, be such that the above equation shall be derivable from them. Hence one of these equations may be rejected. If we reject the first, viz. the one which contains p_1, and then reduce the others by making the λ quantities which are not derived from K to vanish, the system typified by (8.) evidently reduces to the system which we should have to employ if we applied the method of Proposition IV. directly to the system of $n-1$ equations (7.). Hence the quantities $p_2, p_3, \ldots p_n$ satisfy the final conditions to which that application would lead, and therefore by hypothesis the equations (7.) admit of solution by a single system of finite positive values of $y_2, y_3, \ldots y_n$.

Now in general

$$x_i=x_1^{\frac{a_i}{a_1}}y_i.$$

Hence since $x_1=0$ and y_i is finite and positive for all values of i from 2 to n, we see that x_i will be 0 for all values of i for which a_i is positive, finite and positive for all values of i for which a_i is 0, and infinite for all values of i for which a_i is negative.

Case 2; the coefficient a_i negative. Here the inequation of condition (3.) must be supposed to determine the lowest of the superior limits of p_1, and therefore when p_1 coincides with that limit we have

$$a_1p_1+a_2p_2\ldots+a_np_n+b=0.$$

The transformations remaining formally the same as before, the following results will present themselves.

The terms in H and in H_2, $H_3\ldots H_n$ will be affected with negative instead of positive powers of x_1. Hence the same determination of y_2, $y_3\ldots y_n$ from the last $n-1$ equations of (6.), which before followed from the assumption $x_1=0$, will now follow from the assumption $x_1=\infty$, which at the same time satisfies the first equation of (6.).

The equation

$$x_i=x_1^{\frac{a_i}{a_1}}y_i$$

shows, since a_1 is here negative and x_1 infinite, that x_i will be infinite for those values of i for which a_i is negative, finite for those values of i for which a_i is 0, nothing for those values of i for which a_i is positive.

In all these cases the values 0 and ∞ appear as limits of finite positive values. This results from the connexion of the second member of the first equation of the system (6.) with the condition (3.).

Lastly, as the incompleteness of form of V only causes certain terms of the developed determinant of Proposition II. to vanish, but leaves the signs of the terms which remain positive, it follows that as x_1 varies from 0 to infinity, x_2, x_3, $\ldots x_n$ being always determined by the last $n-1$ equations of (1.), the function $\frac{V_1}{V}$ will vary by continuous increase between the limits above investigated, viz. from the highest inferior to the lowest superior limit of p_1. Once, therefore, in its progress it becomes equal to p_1, and all the equations are satisfied together.

The above reasoning establishes rigorously that if the proposition is true for $n-1$ variables, it is true for n variables. It remains then to consider the limiting case of $n=1$.

Here, however, only the complete form of V, viz. $V=ax+b$, leads to a definite value of x, and this, as has been seen, is finite and positive. If we give to V the particular form ax, the equation $\frac{V_x}{V}=p$ becomes

$$\frac{ax}{ax}=p, \text{ or } p=1,$$

which determines p, but leaves x indefinite. If we employ the other particular form $V=b$, we obtain no equation whatever, and here again x is indefinite. But as the

reducing transformations are all definite, the above indefinite forms cannot present themselves in the last stage of the problem when the original equations are independent and admit of definite solution.

The proposition is therefore established.

Application.

The general system of algebraic equations upon which the solution of questions in the theory of probabilities depends, is a particular case of that discussed in Proposition V. Its peculiarity is, that all the coefficients which appear in the function V are equal to unity.

The conditions of possible experience, as determined by the method, agree with the conditions shown in Proposition IV. to be necessary, and in Proposition V. to be sufficient, in order that $x_1, x_2 \ldots x_n$ may be determinable as positive finite quantities. For in both cases the quantities λ_1, λ_2, &c. correspond to the different terms in V, and in both cases the equations among those quantities depend simply on the forms of the functions $V_1, V_2 \ldots V_n$, and therefore ultimately on the form of V, irrespectively of the values of the positive coefficients of V. In both cases the systems of inequations are the same.

It follows, therefore, that precisely when the data represent a possible experience, the probabilities of the ideal events from which in the process of solution the problem is mentally constructed admit of determination as positive proper fractions.

Again, as the process for determining the *à priori* limits of the probability sought rests ultimately upon the assumption that the ratio of any term or partial aggregate of terms in V to V itself is a positive fraction, and as this assumption is satisfied when $x_1, x_2 \ldots x_n$ are positive quantities, it follows that the calculated value of the probability sought will always lie within the limits which it would have had if determined by observation from the same experience as the data.

But though the test last mentioned is one which must necessarily be satisfied by a true method, it is of infinitely less theoretical importance than that from which it is derived, viz. the test which consists in the absolute connexion between possibility in the data and formal consistency in the method.

As the conclusions of Propositions IV. and V. depend upon the form of the function V and the fact that its coefficients are positive, it follows that if in the application of the method to questions of probability we substituted any other positive values for unity in the coefficients of V, leaving the rest of the process as before, we should still be able to determine $x_1, x_2, \ldots x_n$ as positive quantities, or as limits of such, and the altered value of the probability sought would still be consistent with the experience from which the data are supposed to be derived. It would, however, properly speaking, be a value of interpolation, not a probability.

I will close with a few remarks upon the general nature of the method, and of the solutions to which it leads.

1st. The probability determined is not precisely of the same nature as the probabilities given.

For the data are supposed to be derived from experience; and therefore, on the supposition that the future will resemble the past, the events of which the probabilities are given will in the long run recur with a frequency proportioned to their probability.

But the probability determined is always an intellectual rather than a material probability. We cannot affirm that in the long run an event will occur with a frequency proportional to its calculated probability; but we can affirm that it is more likely to occur with this than with any other precise degree of frequency; that if it do not occur with this degree of frequency, the data are in some measure *one-sided.*

At the same time the limits of possible deviation are determined.

2ndly. General solutions obtained by the method do sometimes, but not always, admit* of being verified by other methods. I believe that this is solely because it is not often possible to solve the problem by other methods without introducing hypotheses which are of the nature of additional data, and, in effect, limit the problem. Every general solution, however, admits of a number of particular verifications by necessary consequence from the theorems established in this paper.

3rdly. It has been seen that a calculated probability is not necessarily a definite numerical value. It may be of the form $A+cC$, in which c is an arbitrary positive fraction. Here it is implied that the probability admits of any value between A and $A+C$. If, further, $A=0$ and $C=1$, it is implied that the probability may have any value between 0 and 1,—is therefore quite indefinite. This would really arise if we applied the method to a case in which the event of which the probability is sought had absolutely no connexion with those of which the probabilities are given.

Hence in the present theory the numerical expression for the probability of an event about which we are totally ignorant is not $\frac{1}{2}$, but c†. Hence, also, when all the probabilities given are measured by $\frac{1}{2}$, it is not to be concluded (upon the ground of *e nihilo nihil*) that the probability sought will also be $\frac{1}{2}$.

4thly. While extending the real power of the theory of probabilities, the method tends in some cases to diminish the apparent value of its results. For all problems in which the data admit of logical expression can be solved by it; but the resulting solutions, founded upon the bare data, may be of an indeterminate character, in place of the determinate results to which ordinary methods, aided by hypotheses not really involved in the data, lead.

This is the case with the problem of the combination of different grounds of belief or opinion. The general solution is indefinite. In two limiting cases, however, it assumes a definite form; one of these, which agrees with the formula generally accepted, representing the extreme cumulative force of testimonies, the other the mean weight of

* Professor DONKIN has verified a general solution (Laws of Thought, p. 362).

† See on this subject a paper by Bishop TERROT, Edinburgh Transactions, vol. xxi. part 3.

judgments. Both these, however, occur as limiting cases, and they can only be applied with confidence under extreme circumstances, such as probably never occur in human affairs. (*Edinburgh Memoir*, pp. 630–645.)

5thly. I have, in effect, remarked that there is reason to suppose that all questions in the theory of probabilities can ultimately be reduced to questions in which the immediate subjects of probability are *logical*, i. e. involve no other essential relations than those of genus and species, whole and part. This is a question of theoretical rather than of practical interest. For instance, whether the formula of the arithmetical mean, which is the basis of the theory of astronomical observations, is self-evident, or whether it rests upon an ultimate logical basis, or whether, as I am inclined to believe, it may lawfully be regarded in either of these distinct but not conflicting lights, the superstructure remains the same.